

Data Wrangling with SQL

Good morning!

We start at 10 : 00



- This is NOT a DB Administration class (but we will do some DB Admin tasks...)
- This is NOT an SQL class (but we will write a lot of SQL)

What this course will teach you

Show we
the list of the Business need
↓ customers will
↓ highlight accounts

Go to the Technical DB of customers
and get the 5 highest

Question → Solution

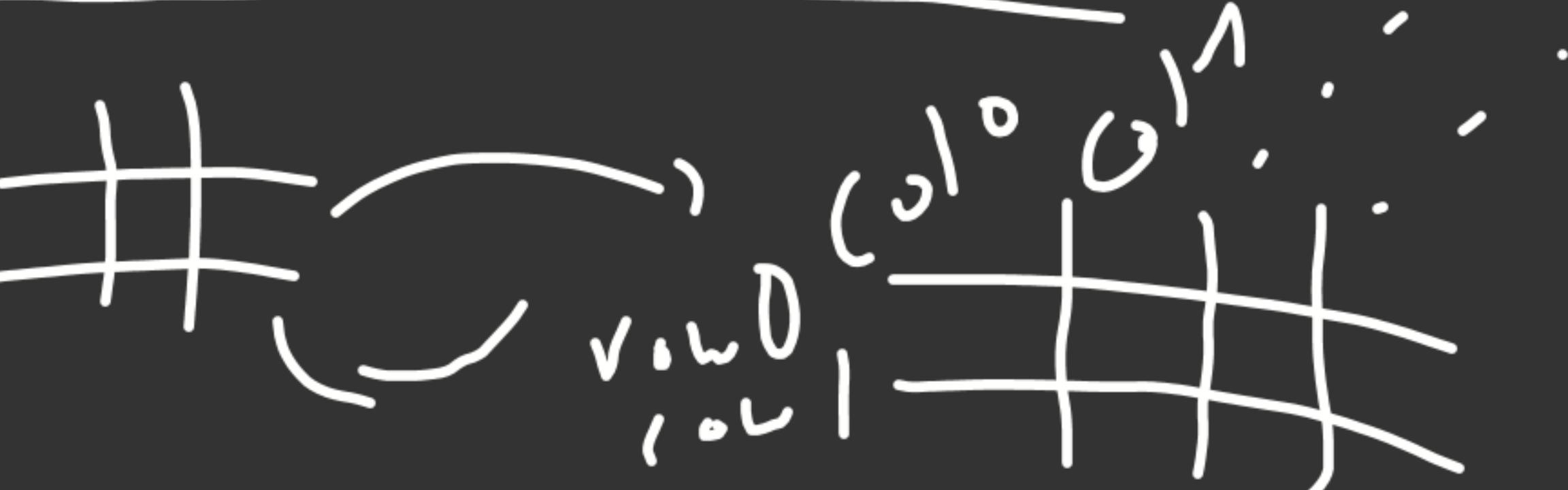
select SQL
select top 5 from
customers

What is a RDBMS()

RDBMS

Relational Database Management System

What is a RDBMS (2)

- Software
 - Stores data in tables
 - Retrieves data by querying
- 
- The diagram illustrates a relational database table structure. It features a grid of four columns and three rows. The first column contains the values 'John', 'Mike', and 'Steve'. The second column contains 'Doe' repeated three times. The third column contains '100' repeated three times. The fourth column contains '1000' repeated three times. A large hash symbol (#) is positioned above the first column, and a circled 'View' label is placed above the second column.

++ security Some RDBMS *-- security*

Proprietary \$

- MySQL

- Oracle

- DB/2 (IBM)

- Hadoop

- AWS Aurora

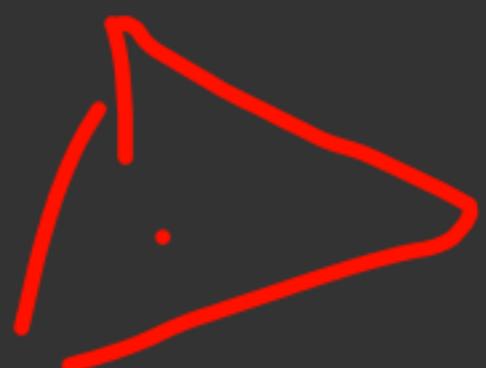
Open-source Ø

- MySQL

- PostgreSQL

- MariaDB

- SQLite



RDBMS Architecture

Client

= Client - Server
Architecture



- UI (e.g., webpage)
- Software (e.g., SQL Management Tools, Specifier)
- API (e.g., Python, SQLite)

Server

DB engine

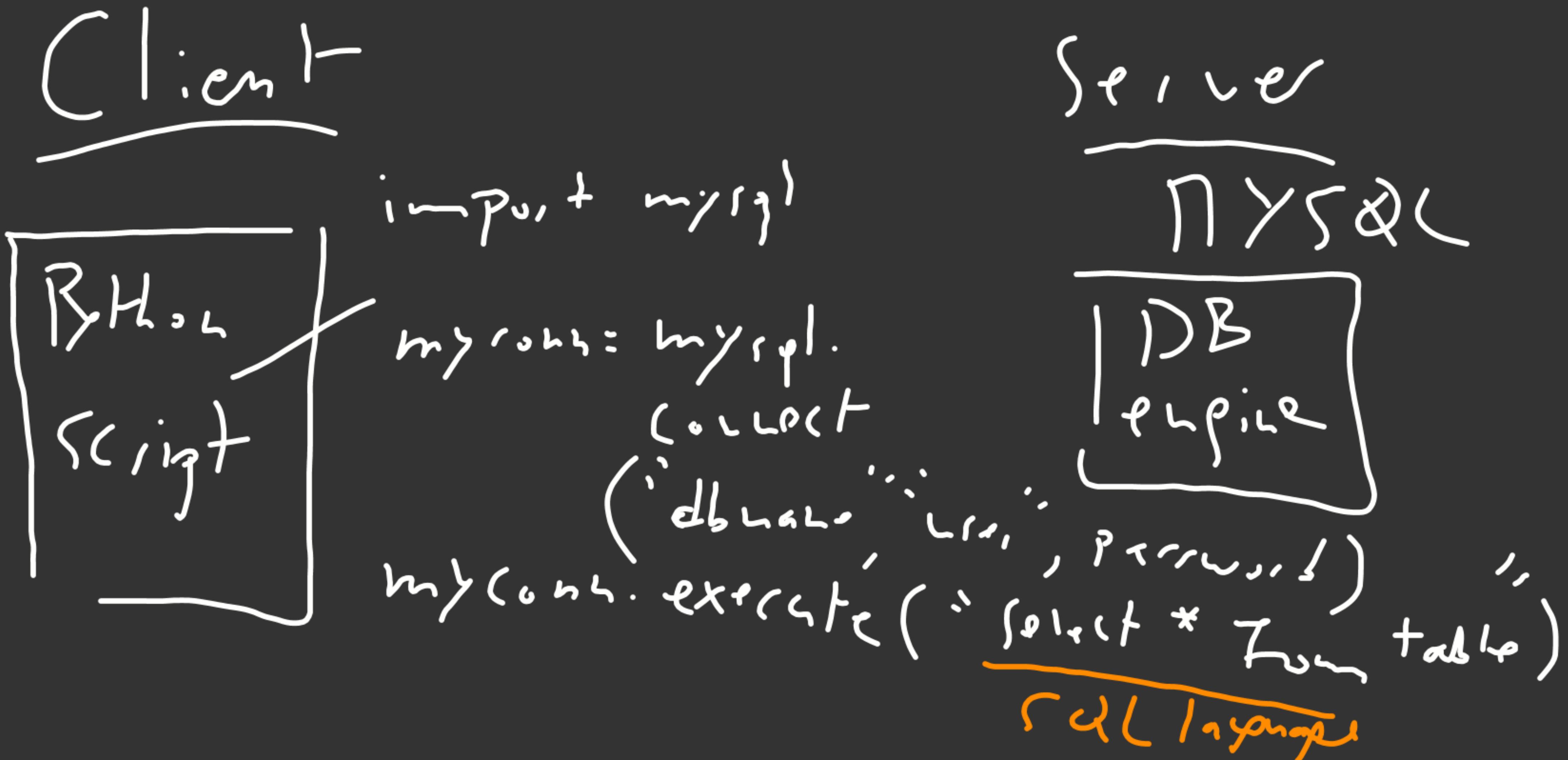
TSSQL

Storage / Implementation
query interpreter

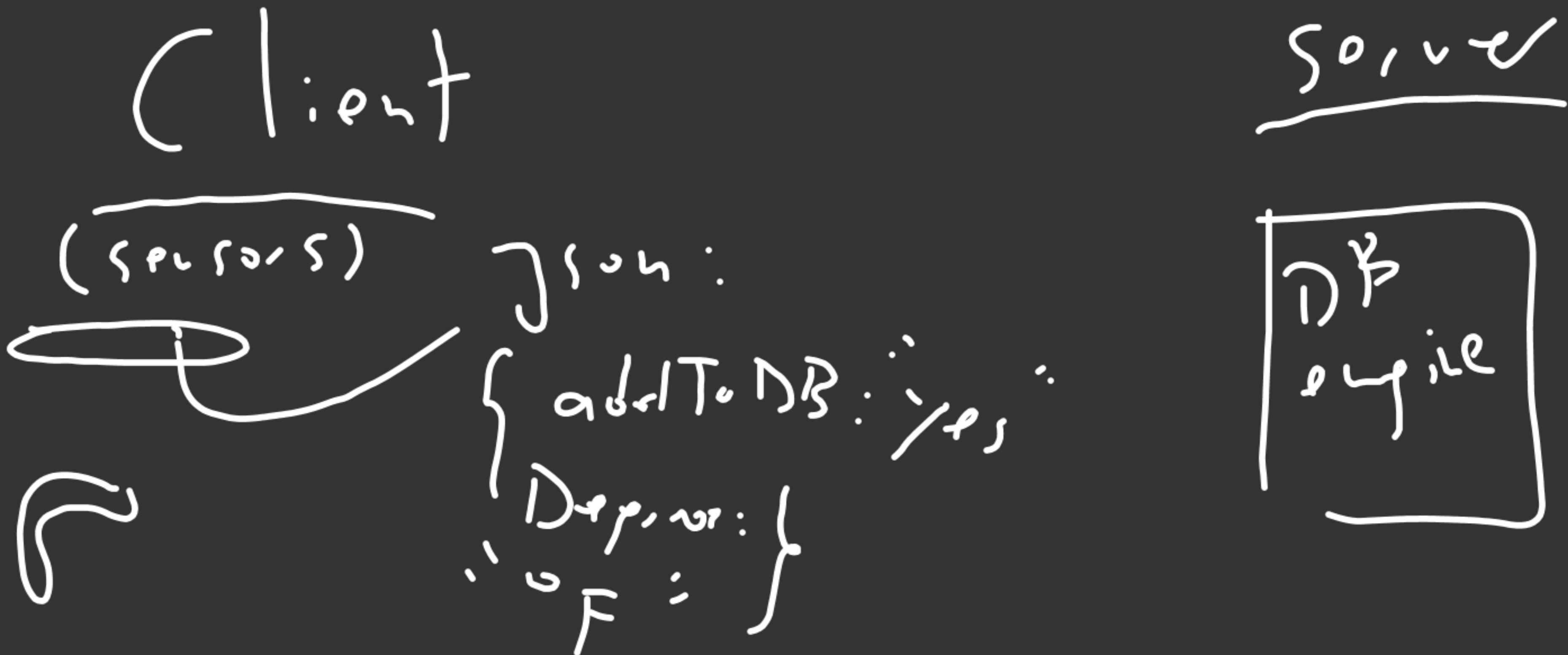
UI example

CUSTo.NERS					Column = information
email	password	first name	last name	high	1 row
...	1 AUS profile
...	(1 user)

API example (1)

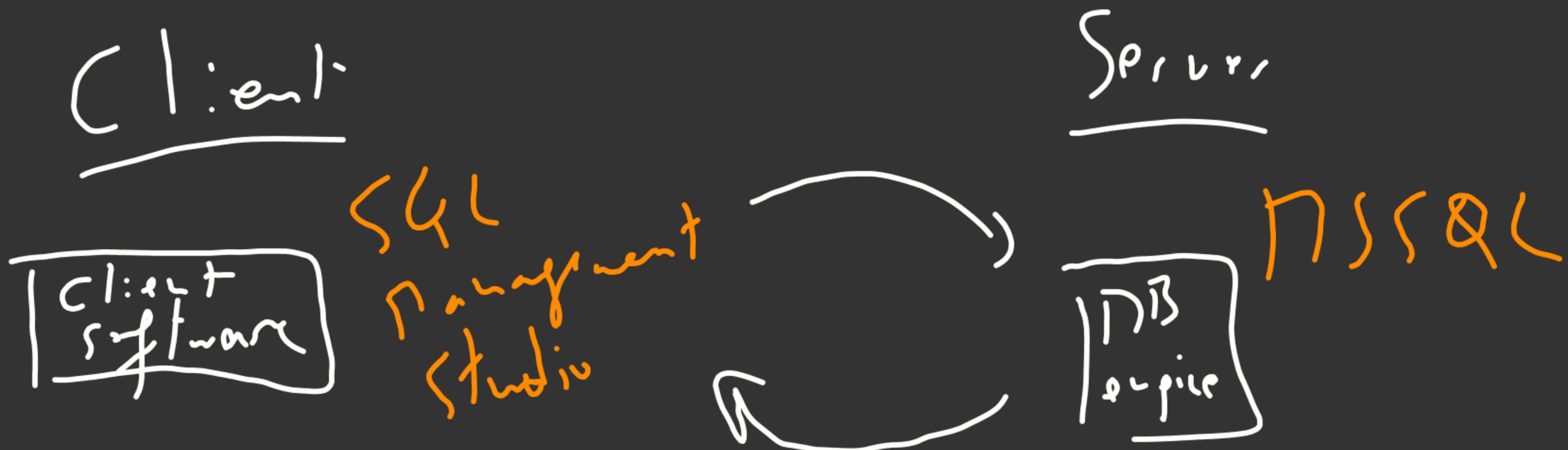


API example (2)



Software example

→ Our case in this course !



Break

Back at 11 : 30

Installation Options (I)

- Local
- VM
- Hybrid

These

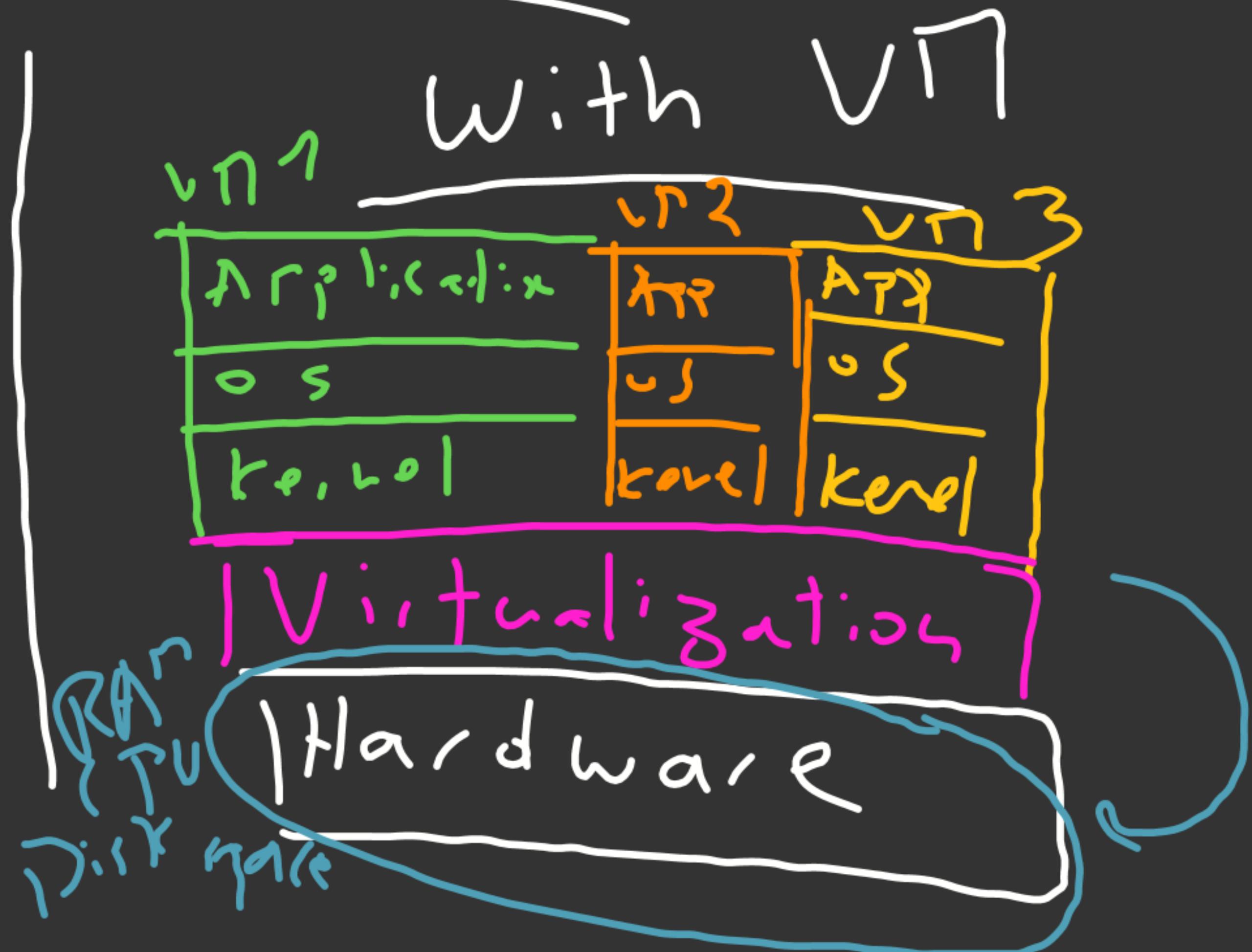
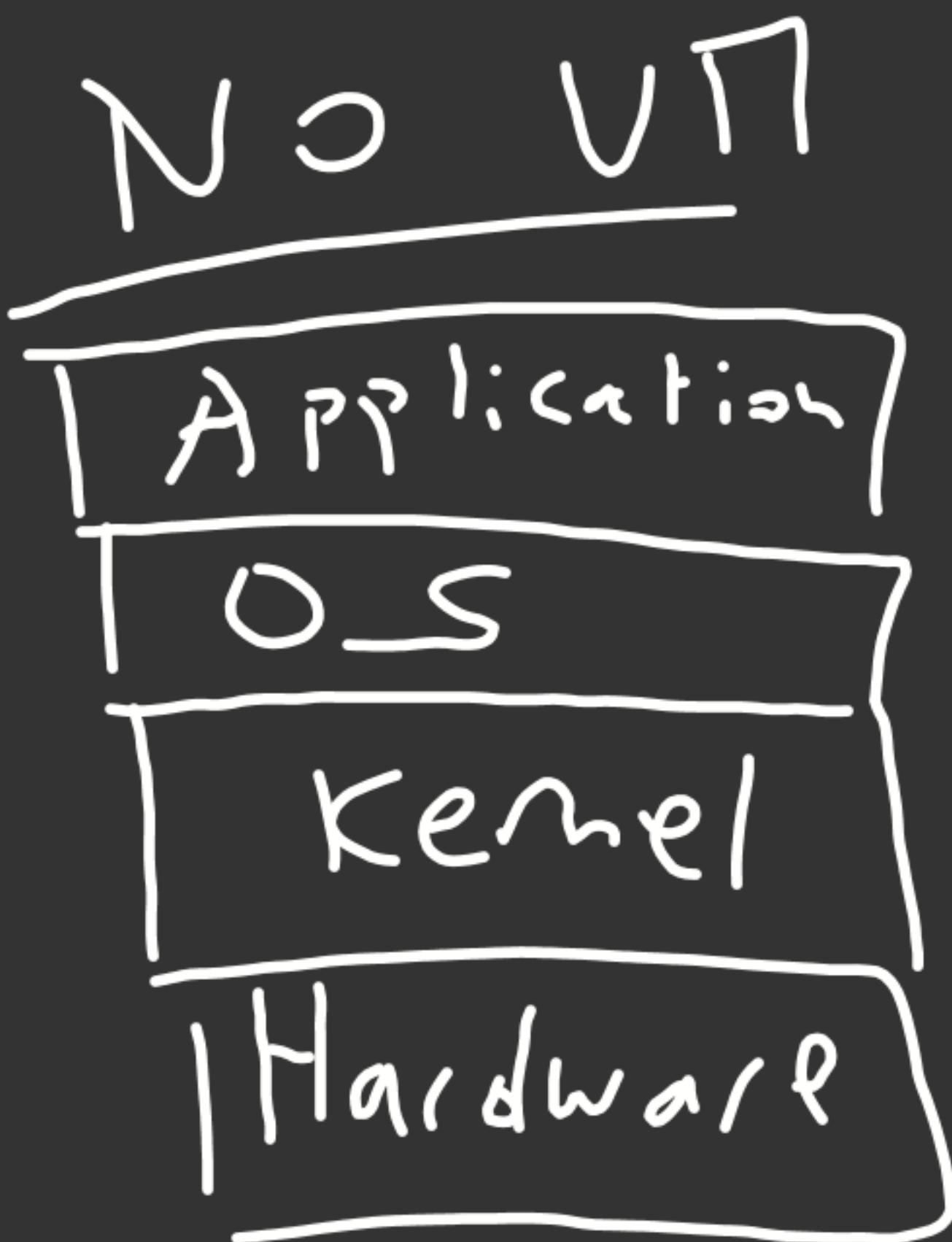
Client
Server
Architecture

A note on the exam

- 2 attempts
 - 5 hrs/attempt (10 hrs in total)
 - 5 questions
 - Self-graded
- Module

Q1	True 20%
X False	0%
Q2	

Virtualization



Installation options (2)

- Local Installation → skip
and install DB directly
- VM → Prepare a VM

Lunch

Back at 13:45

Local VN Installation

- ① Create & configure VN ✓
- ② Install an OS (windows / linux) ✓
- ③ Install RDNS (server + client)

Break

Back at 14:45-

MySQL Installation (1)

A screenshot of a Moodle course page. On the left, there's a sidebar with a background image of a field. The main content area lists several items:

- Data Wrangling with SQL Syllabus
- Biography Hanna Abi Akl
- Python for Data Analysis Data Wrangling with Pandas, NumPy, and IPython by Wes McKinney
- [IMPORTANT] Microsoft Azure Education - Software Download
- en_sql_server_2019_developer_x64_dvd_baea4195** (circled in orange)
- SQL Server Management Studio (SSMS) - Download
- WideWorldImporters Database Dump
- RelationalModel Introduction
- Individual Mini Project - Data Wrangling with SQL

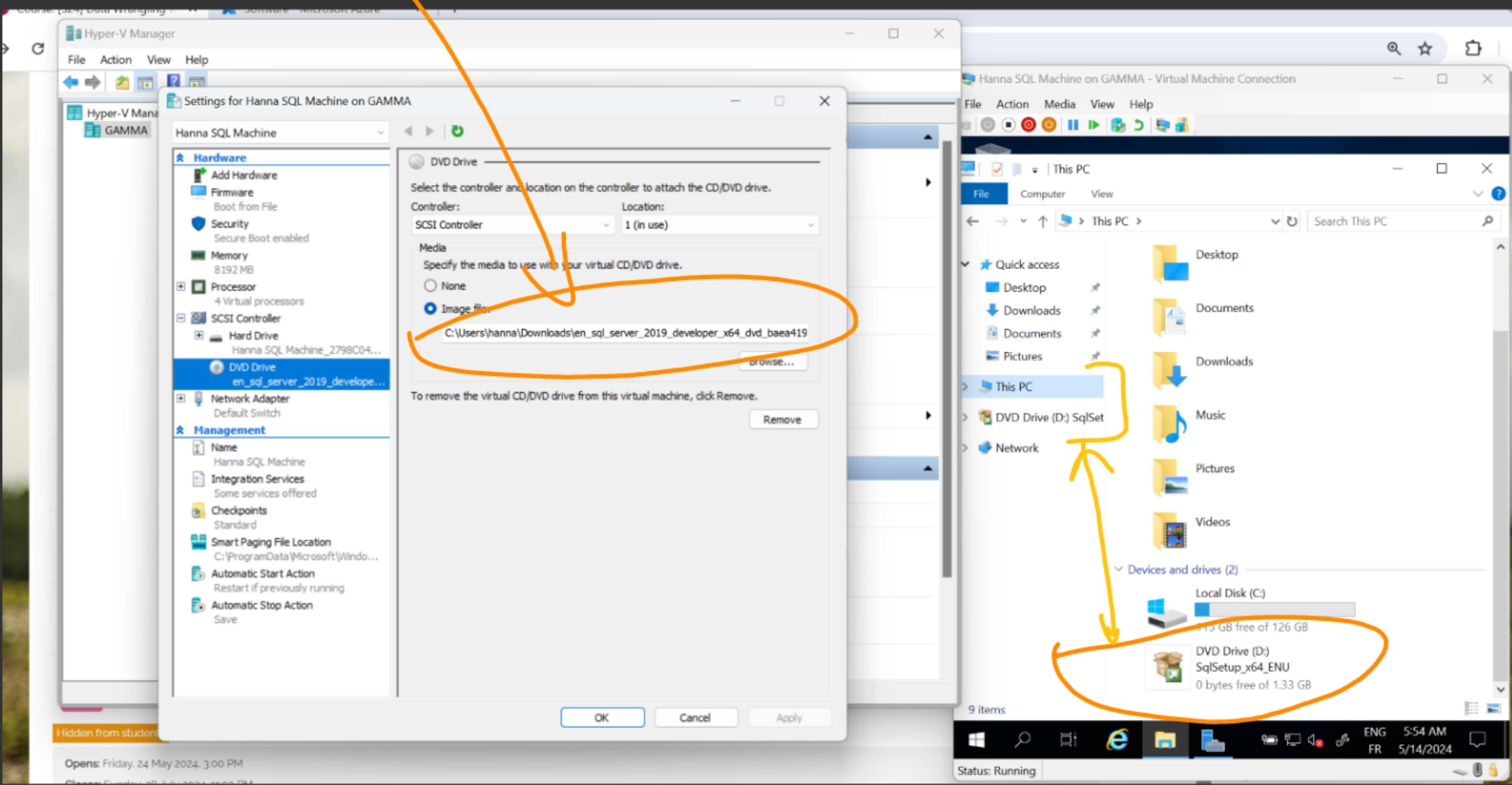
An orange arrow points from the circled link to the "Download" menu in the top right corner. Handwritten orange text "Download this" is written next to the arrow.

The top right corner shows a navigation menu with the following items:

- Users
- Unenrol me from this course
- Filters
- Reports
- Gradebook setup
- Import
- Backup
- Restore
- Copy course
- Reset
- Question bank
- Accessibility toolkit
- Recycle bin
- Download course content
- Site administration

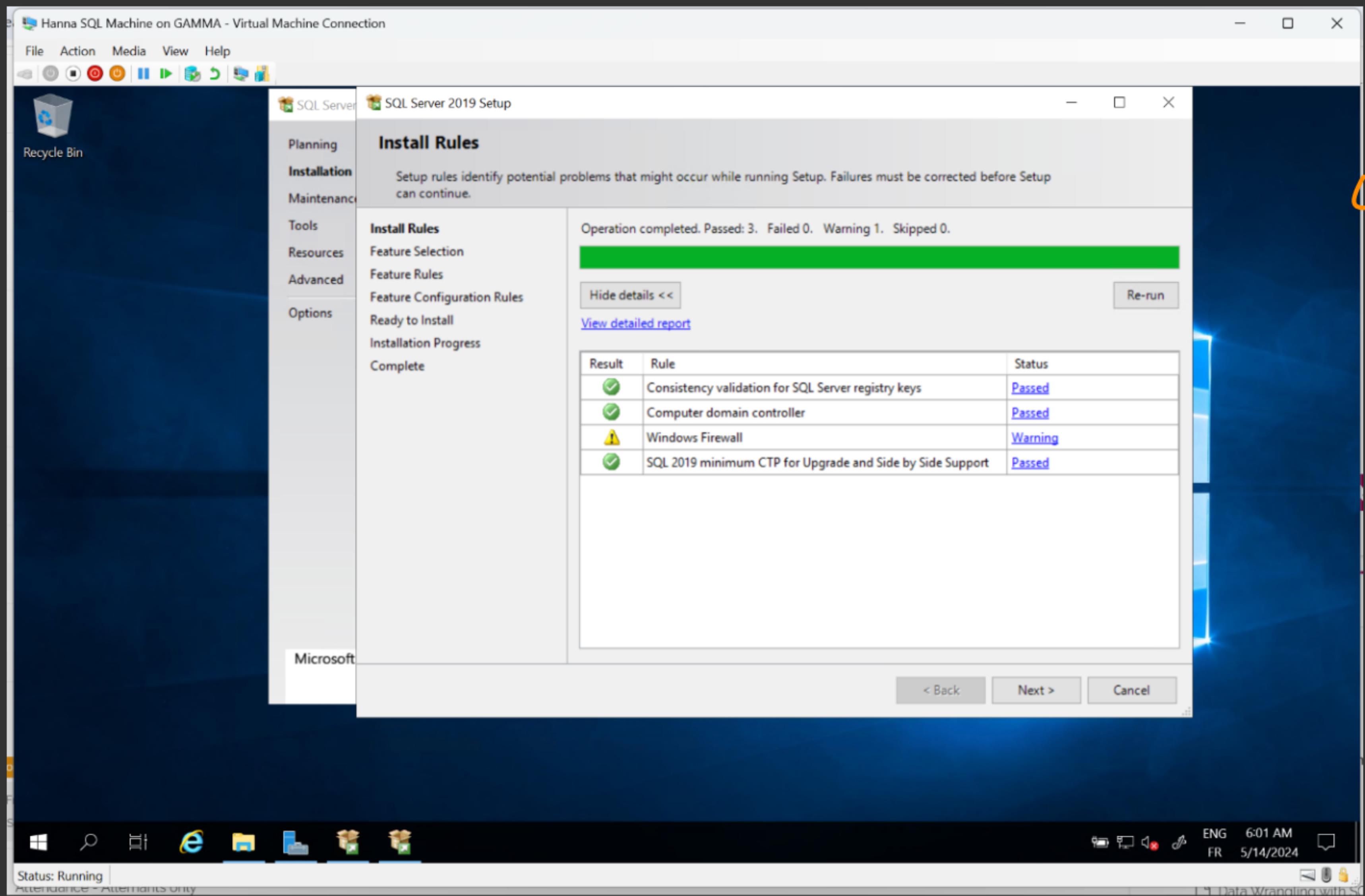
Below the menu is a search bar with the placeholder "Search in settings" and a magnifying glass icon. At the bottom is a "Navigation" sidebar with links to Dashboard, Site home, and Site pages.

① Download from Noodle MSSQL Installation (2)

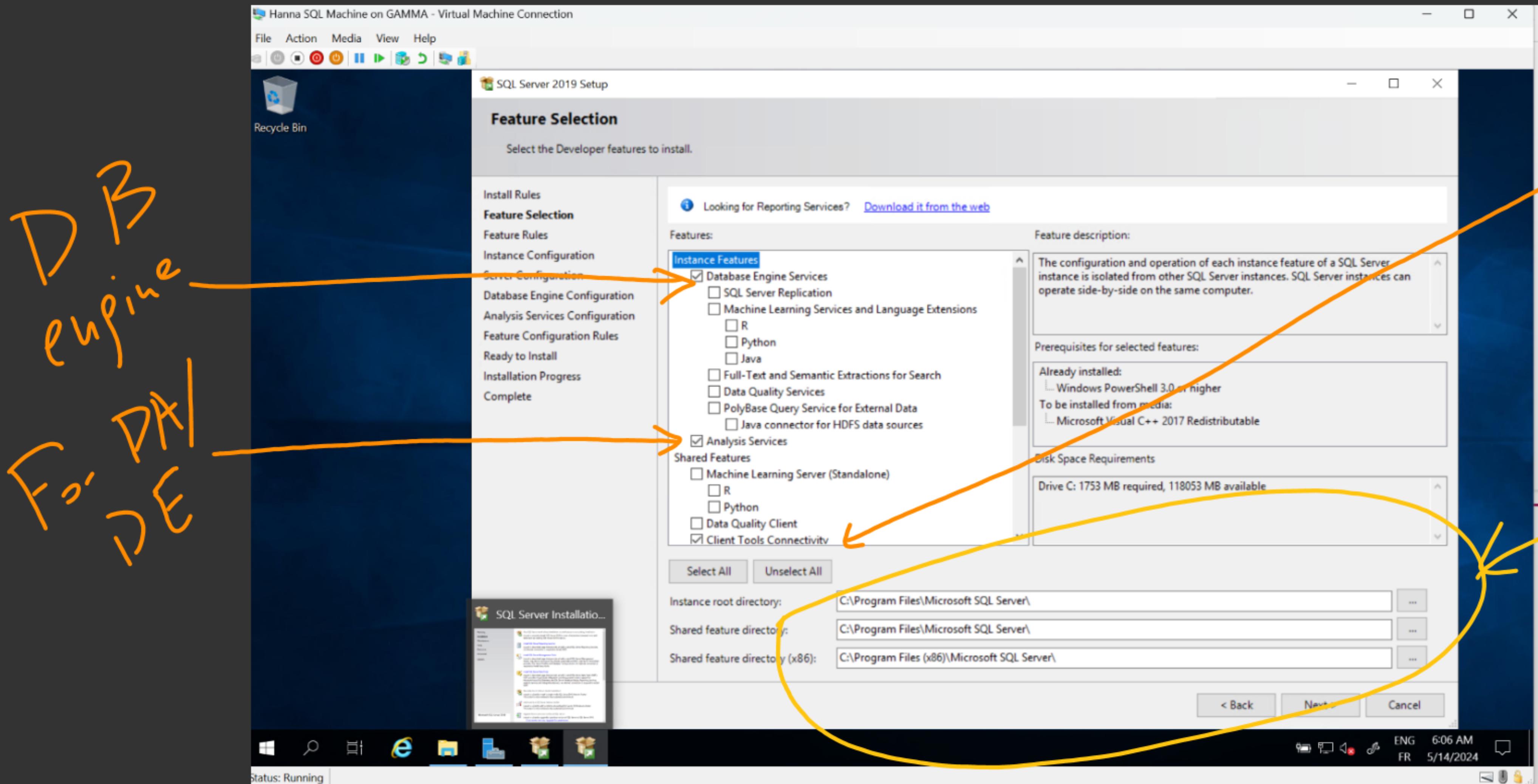


② Start VTM and find the image

MySQL Installation (3)

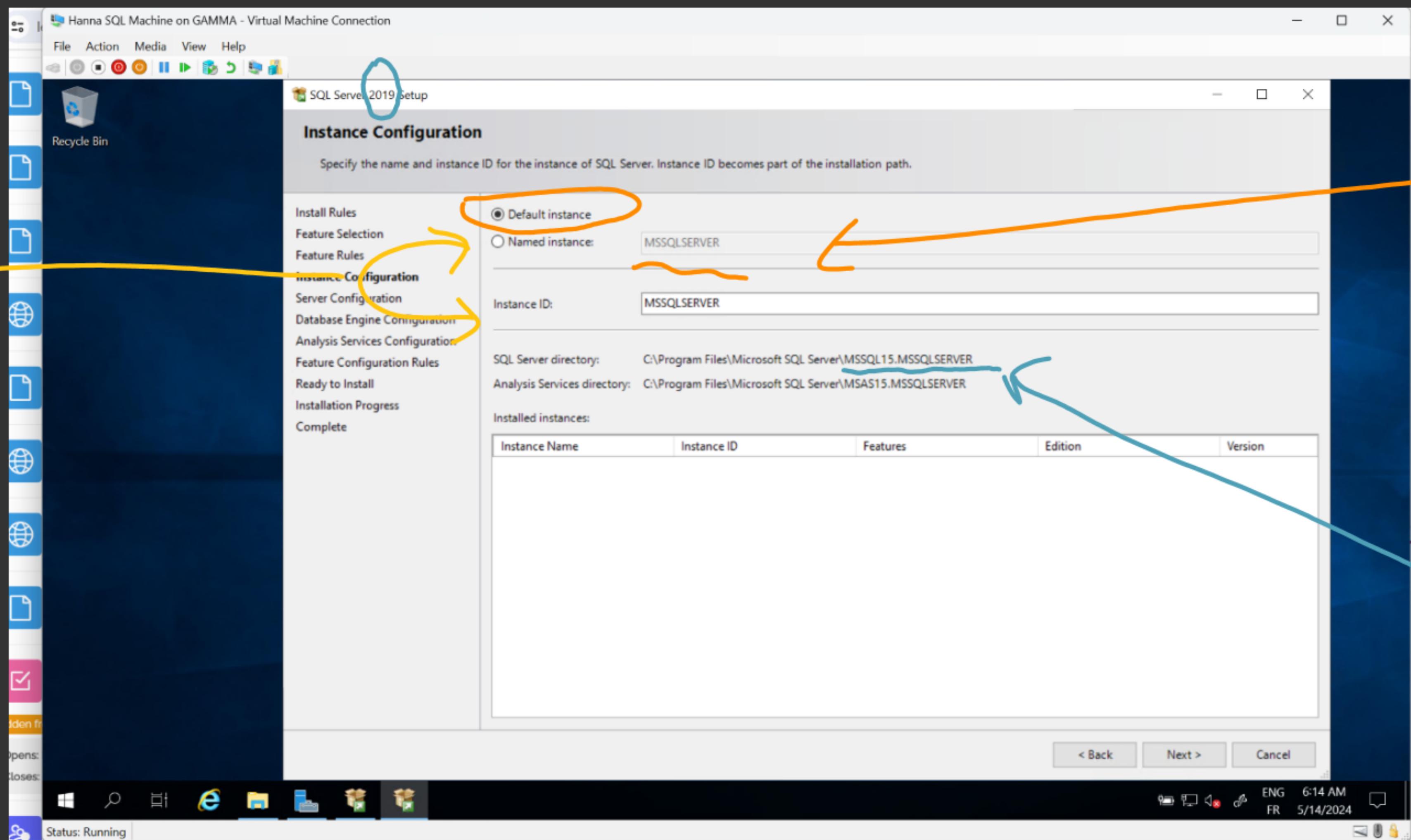


MySQL Installation (4)



SQL Installation (5)

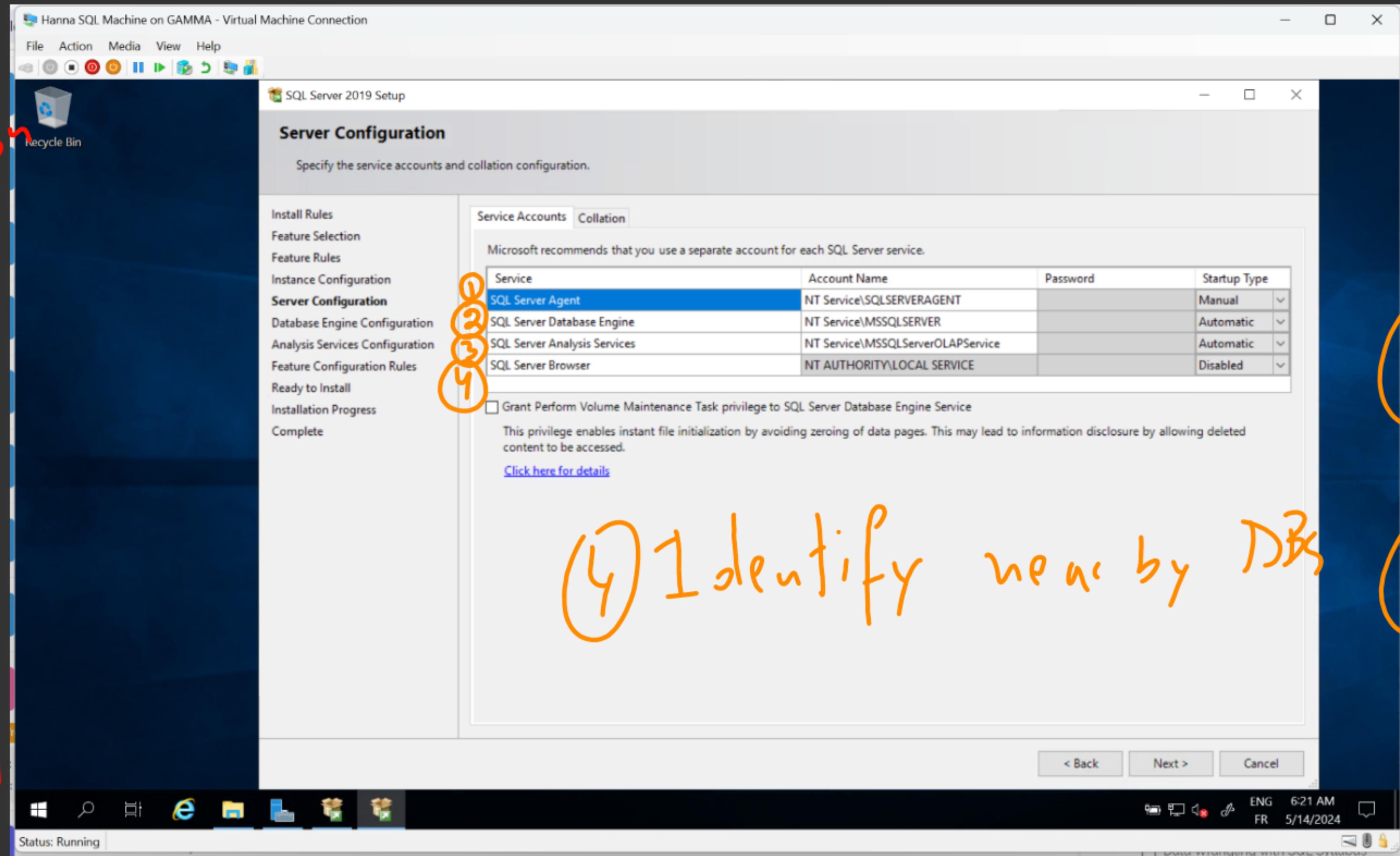
Setup
if you
want
multiple
instances



Default
instance
Named
build
name
version

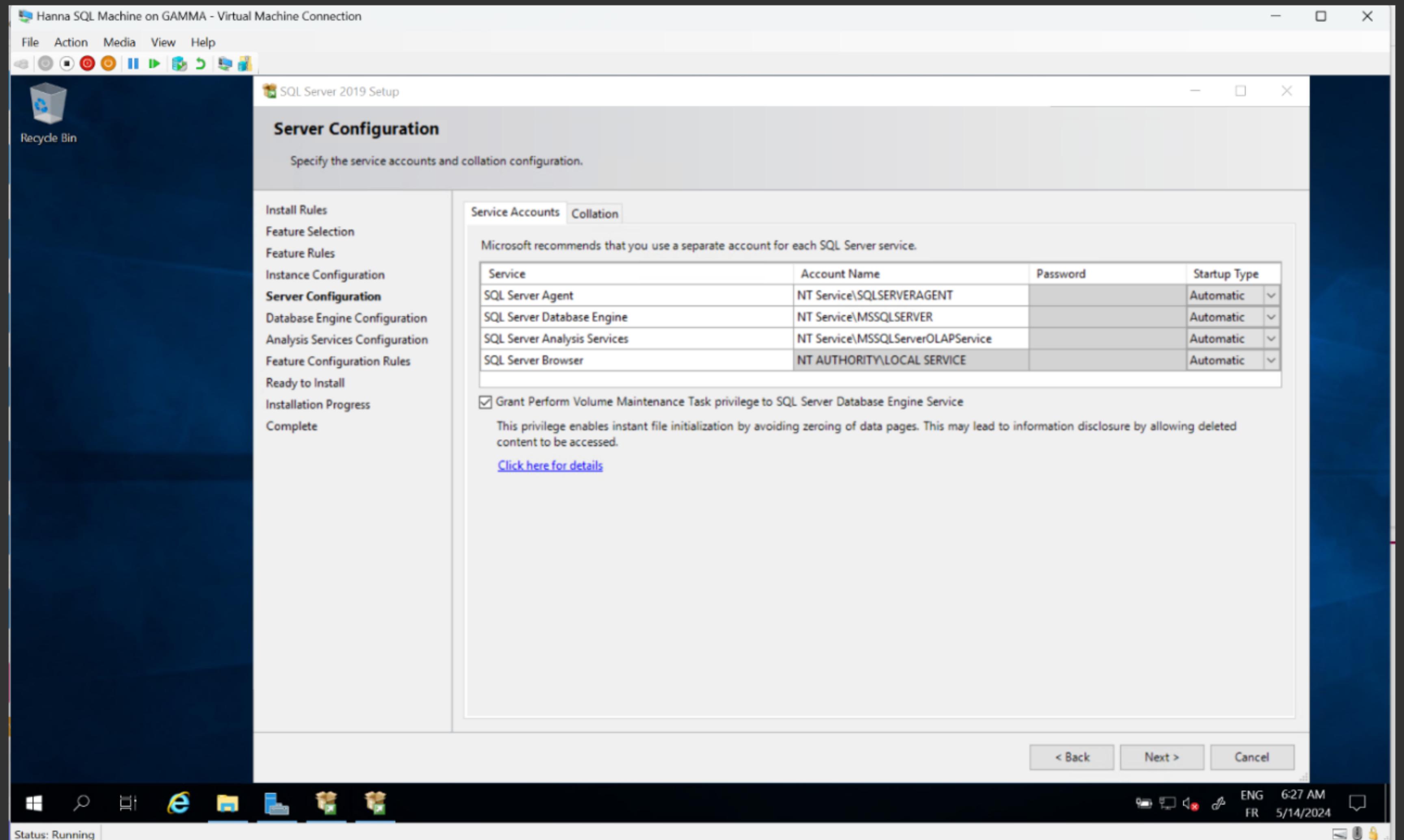
MSSQL Installation (6)

This configuration is wrong
(go to next slide for correct one)



- ① Starting DB service
- ② DB engine
- ③ Analytics engine

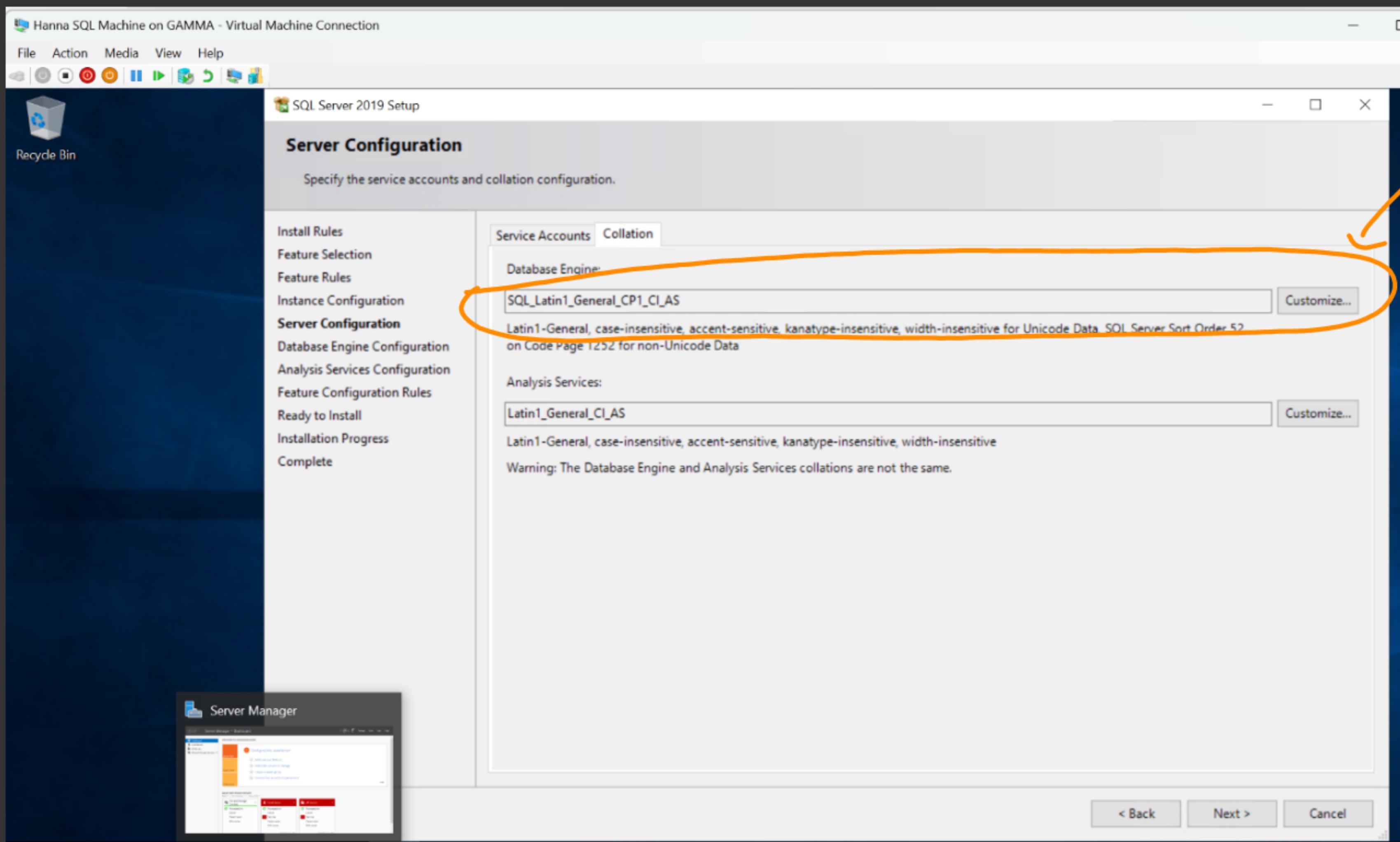
MySQL Installation (?)



Correct
Configuration



SQL Installation (8)



Tells
DB
how
read
of
interact
charsets

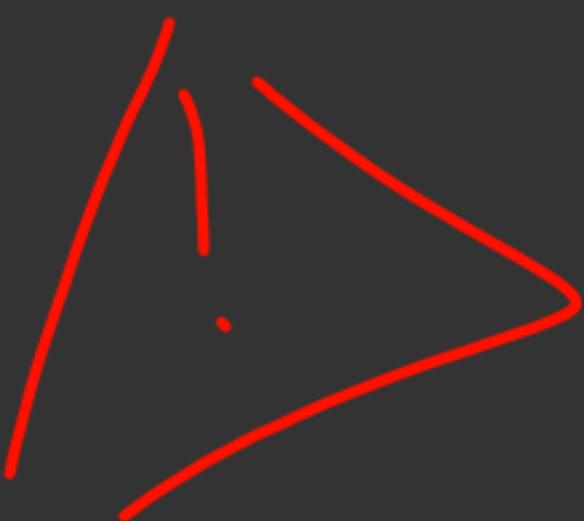
$\acute{e} = e ?$ $\underline{DB \text{ Collation (1)}}$ 
 $L = B ?$

Select * from customers where FName = "Sébastien"

Select * from customers where FName = "Sebastien"

Select * from customers where FName = "SEBASTIEN"

Select * from customers where FName = "SeBastIen"



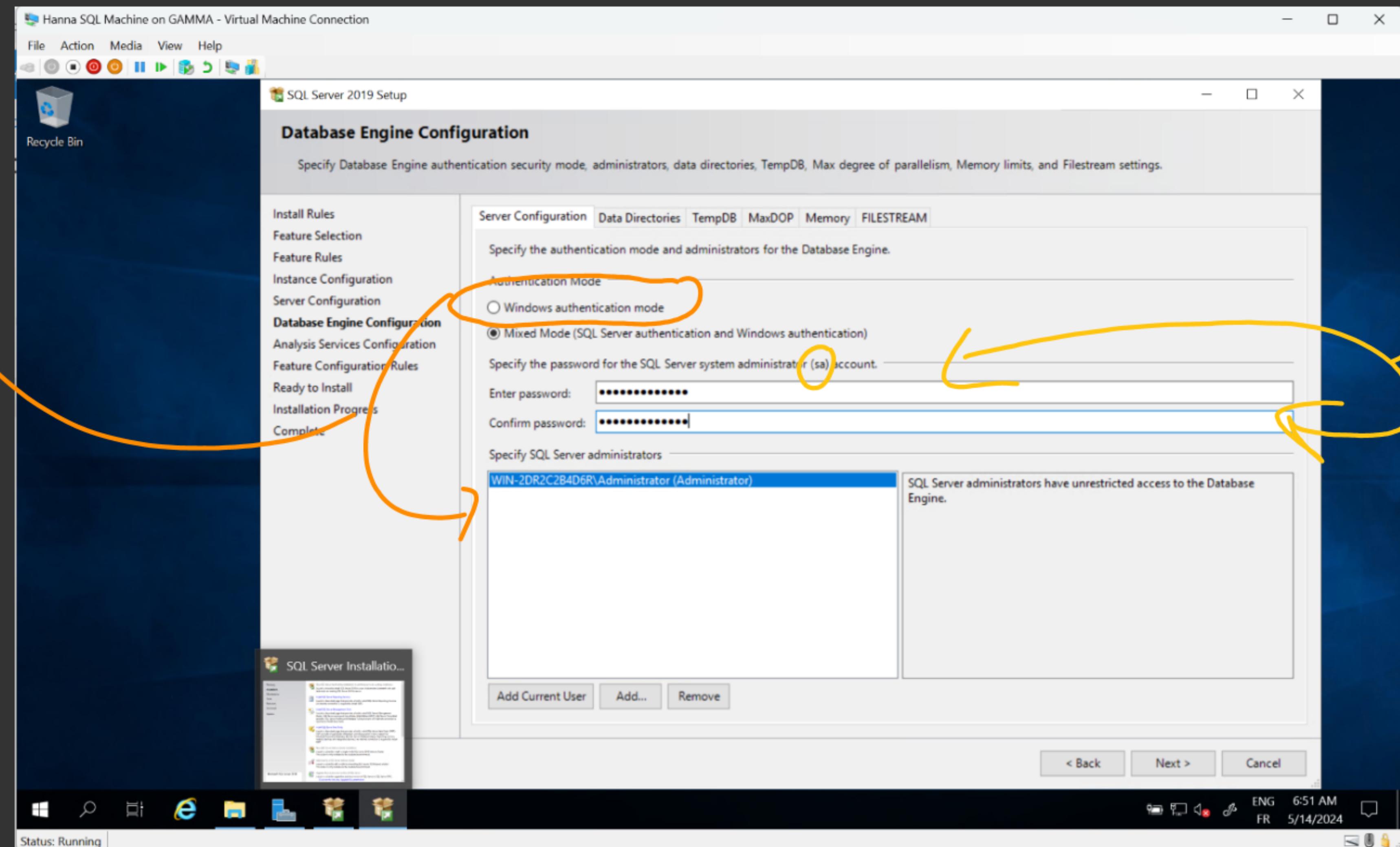
DB Collation (2)

- Collation tells DB how to treat characters
- It is very hard to change once it is set
- Default: SQL_LATIN_CI_AS
 - sebastien = Sébastien = seLASTien = Sébastien
 - Sébastien ≠ sebastien

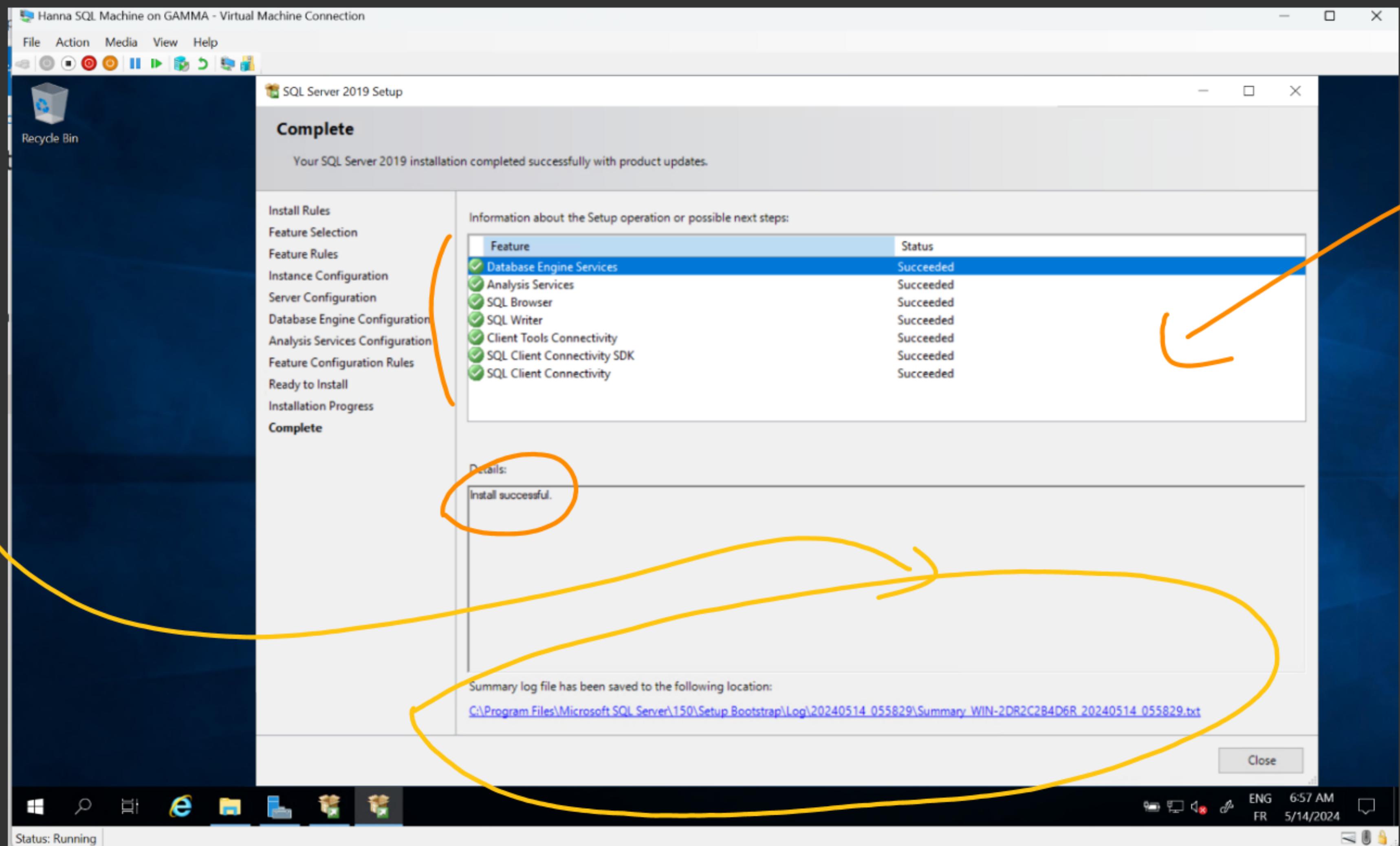
MySQL Installation (9)

Vn|local
Vhel

Dedicated
MySQL
User



If install fails, open a ticket and add this file



This means installation succeeded!!

Break

Back at 16:20

SSMS Installation

Download
from
Noolle

The screenshot shows a Moodle course page with a sidebar on the right. The main content area lists several download links:

- Data Wrangling with SQL Syllabus
- Biography Hanna Abi Akl
- Python for Data Analysis Data Wrangling with Pandas, NumPy, and IPython by Wes McKinney
- [IMPORTANT] Microsoft Azure Education - Software Download
- en_sql_server_2019_developer_x64_dvd_baea4195
- SQL Server Management Studio (SSMS) - Download** (this link is circled in orange)
- WideWorldImporters Database Dump
- RelationalModel Introduction
- Individual Mini Project - Data Wrangling with SQL

The sidebar contains the following navigation links:

- Users
- Unenrol me from this course
- Filters
- Reports
- Gradebook setup
- Import
- Backup
- Restore
- Copy course
- Reset
- Question bank
- Accessibility toolkit
- Recycle bin
- Download course content
- Site administration

Below the sidebar is a search bar with the placeholder "Search in settings" and a magnifying glass icon.

The bottom of the sidebar shows the "Navigation" section with links to "Dashboard", "Site home", and "Site pages".

DB Import()

Download DB

dump

The screenshot shows a Moodle course settings page. On the left, there is a list of files:

- Data Wrangling with SQL Syllabus
- Biography Hanna Abi Akl
- Python for Data Analysis Data Wrangling with Pandas, NumPy, and IPython by Wes McKinney
- [IMPORTANT] Microsoft Azure Education - Software Download
- en_sql_server_2019_developer_x64_dvd_baea4195
- SQL Server Management Studio (SSMS) - Download
- WideWorldImporters Database Dump** (this file is circled in orange)
- RelationalModel Introduction
- Individual Mini Dataset - Data Wrangling with CO

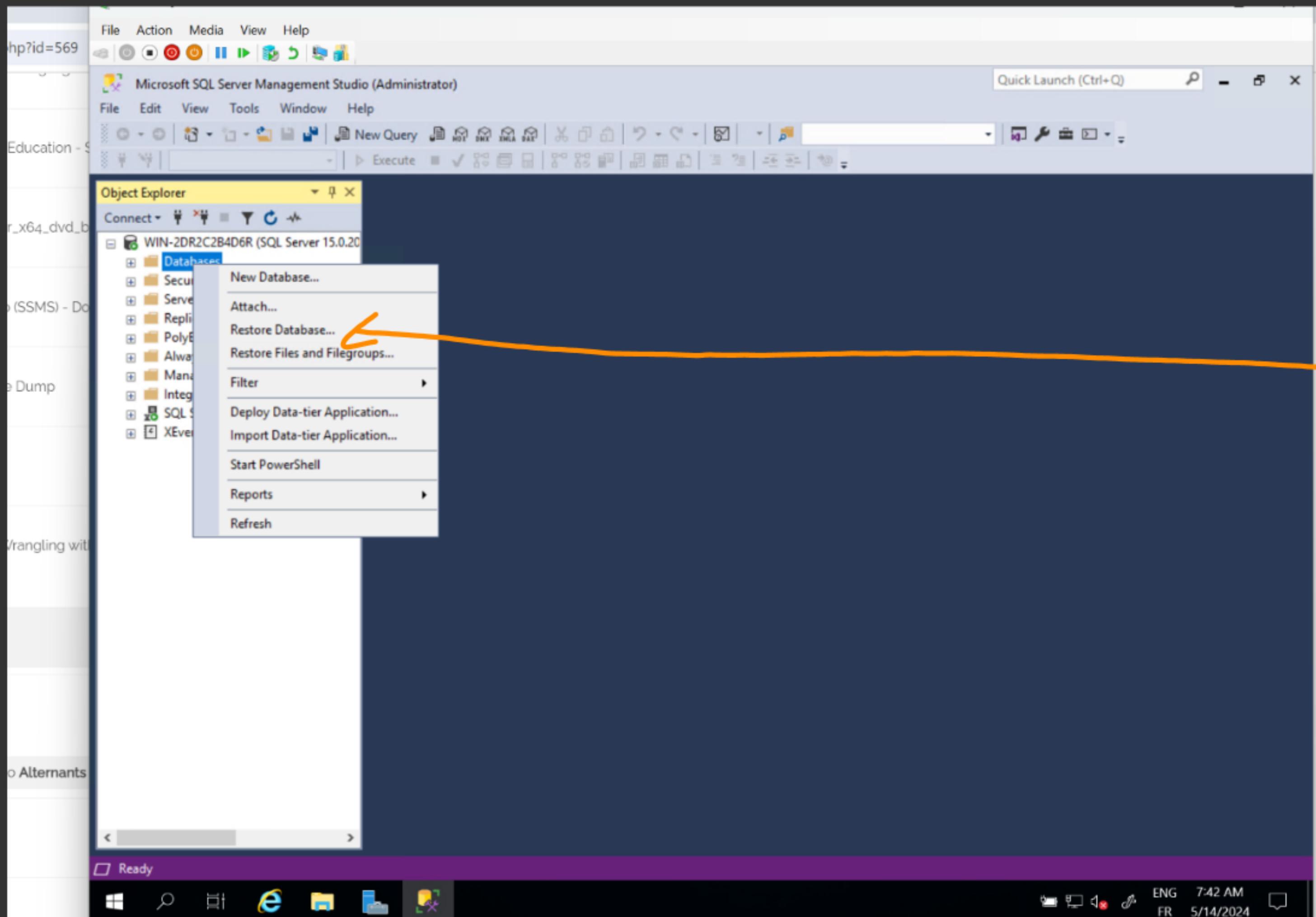
On the right, there is a sidebar menu with the following items:

- Users
- Unenrol me from this course
- Filters
- Reports
- Gradebook setup
- Import
- Backup
- Restore
- Copy course
- Reset
- Question bank
- Accessibility toolkit
- Recycle bin
- Download course content
- Site administration

Below the sidebar is a search bar labeled "Search in settings" with a magnifying glass icon.

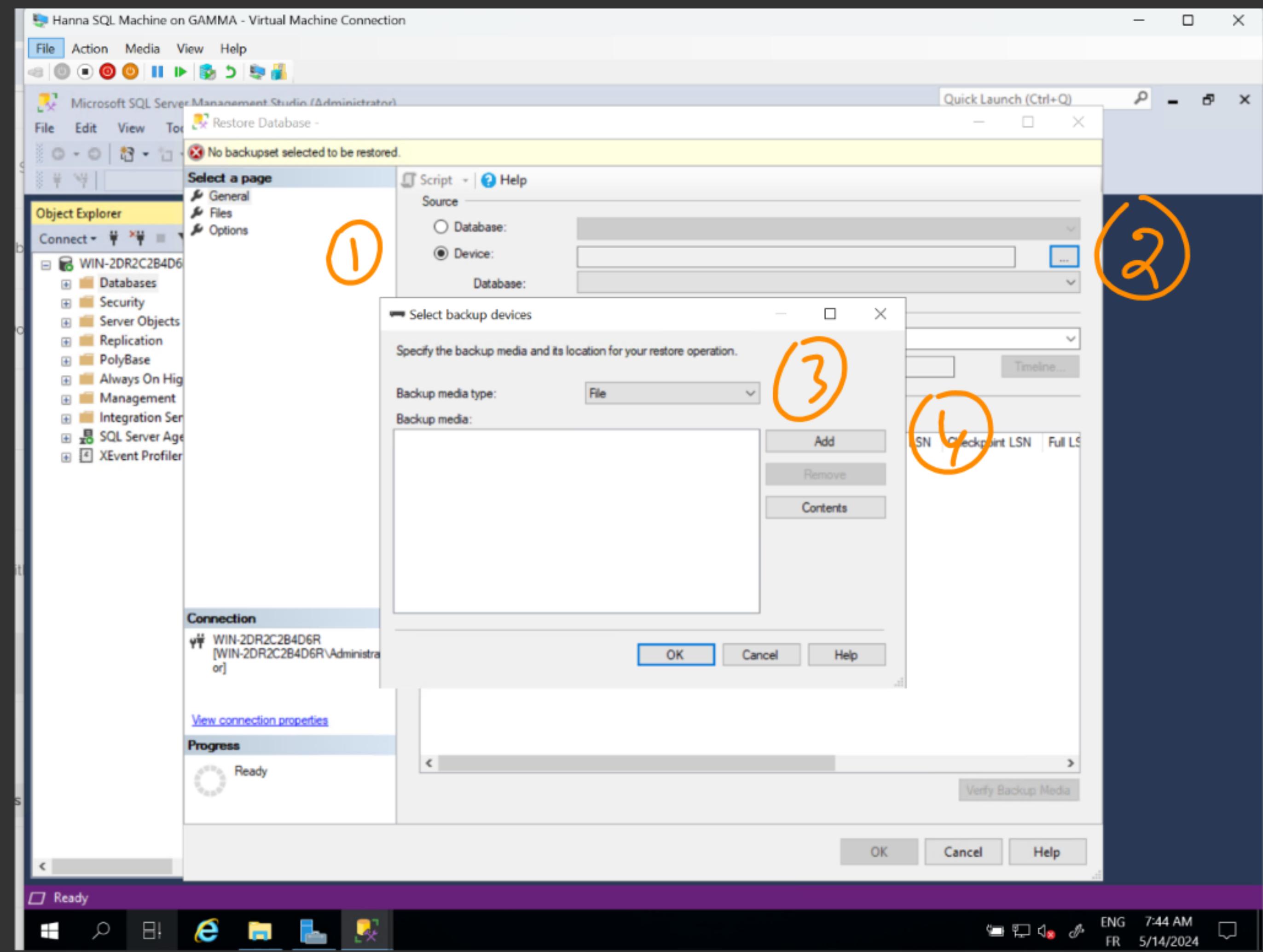
At the bottom of the sidebar, there is a "Navigation" section with links to "Dashboard", "Site home", and "Site pages".

DB Import (2)



Click
→
Restore
Database

DB Import (3)

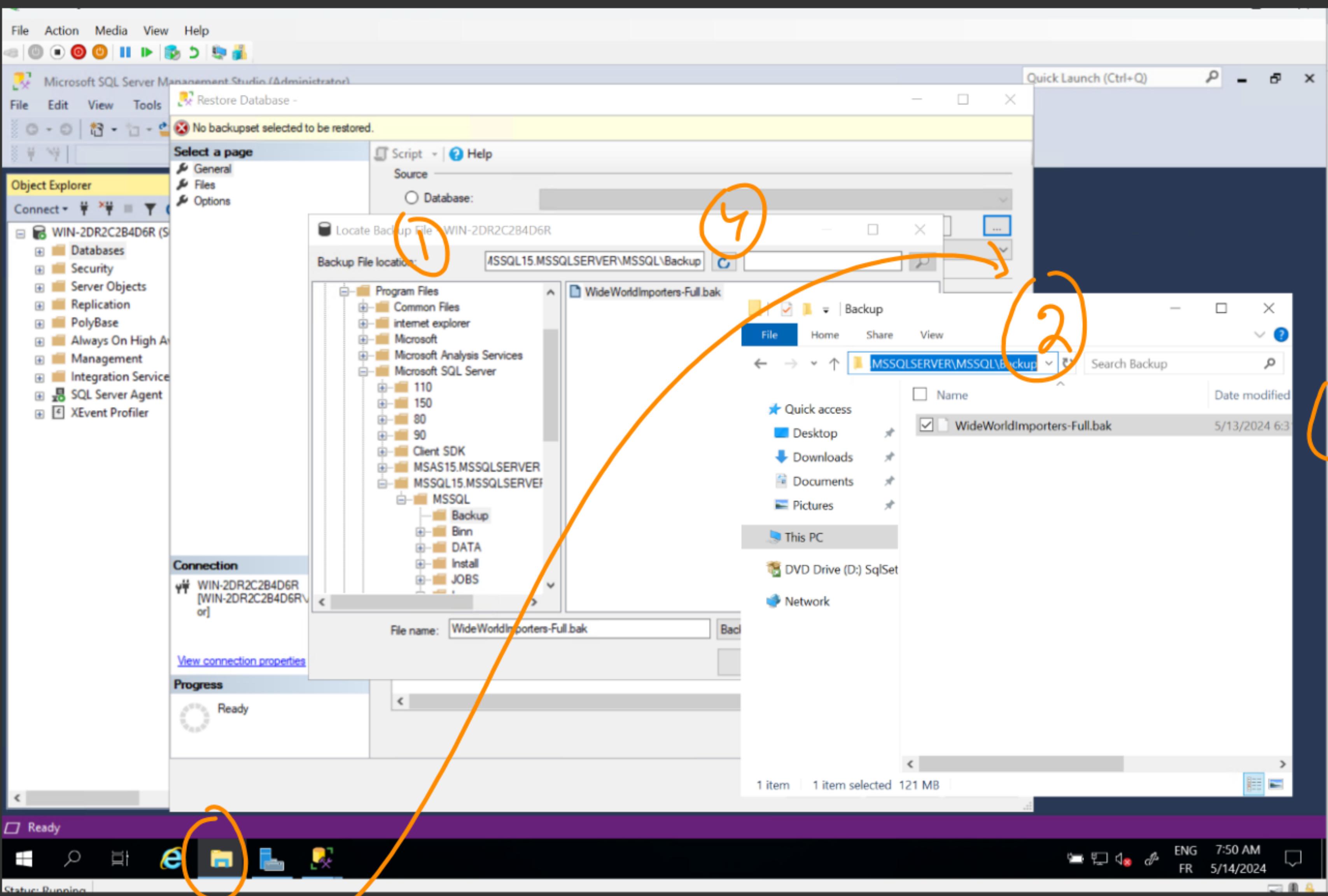


① Copy Path

② Paste Path

③ Put backup dump in path location

DB Import (≤)



④ Refresh until the dump file appears

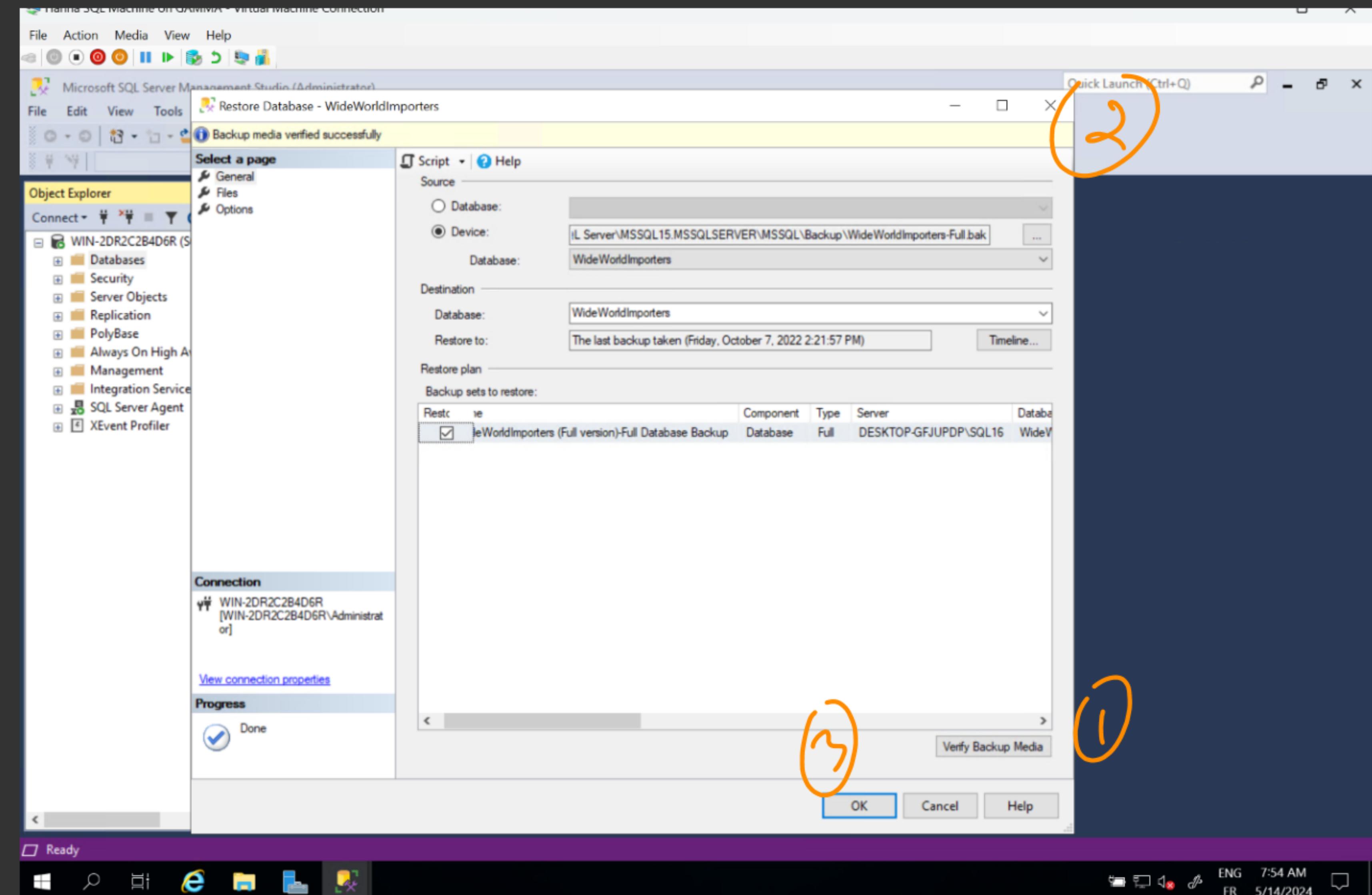
DB Import (5)

① Click

> verify
Backup
Pending

② Make
sure

the banner
appeals



③

Click
ok

①

②

③